


# Colloquialization as a key factor in historical changes of rational and emotional words

Kun Sun<sup>a,1</sup> 

Scheffer et al. (1) argue that in today's "post-truth era" there has been a drastic change from fact-based argumentation to emotion-laden language (i.e., changes of rational/emotional words), a trend paralleled by a shift from collectivistic to individualistic language. The finding is exciting but unconvincing. Their study takes a "culturomics" approach to investigating human behavior and cultural trends through a quantitative analysis of digitized texts (2). Scheffer et al. posit that the surge in post-truth discourse began around the 1970s. Yet such lexical frequency changes do not necessarily indicate a shift from collectivistic to individualistic language. We offer a more plausible way of analyzing these changes.

The distinction between rational and emotional words can be seen as roughly equivalent to the difference between formal and informal words. Scheffer et al. (1) claim rational words are related to rationality, science, and quantification, yet these rational words are actually used in formal situations, particularly in professional and academic writing. Emotional words concerned with intuition, believing, and spirituality are more likely used in informal circumstances (e.g., *imagine*→*envisage*, *fear*→*apprehension*). We collected 210 pairs of formal words and their corresponding informal words (most words here are different from those in Scheffer et al. (1); <https://engdic.org/formal-and-informal-words/>). After obtaining their historical frequencies from Google Books corpora (1850 to 2019), we adopted the method used by Scheffer et al., namely, the second principal component (PC2), which reveals a U-curve (see Fig. 1). The result shows that the historical changes in the use of formal words are basically identical to those of rational words. The same holds for informal and emotional words. These changes can be explained from the perspective of stylistic

formality and informality, which, however, is not the same as a shift from collectivism to individualism.

The historical shift towards a lesser degree of formality is in line with the lexical "colloquialization" trend which has been found in recent numerous studies concerning language changes. Refs. 4–6 proposed that a significant stylistic shift in 20th-century English is due the way in which written language has become more similar to spoken language and more tolerant of various degrees of informality (i.e., colloquialization). Moreover, ref. 1 also reported the historical changes in frequencies of personal pronouns, and these changes are connected with the trend of using passive constructions (3) because the use of the passive voice affects the frequency of personal pronouns. Frequency changes of passive constructions, as one inverse metric of colloquialization, can somewhat explain the changes of personal pronouns. The trend toward colloquialization can more reasonably explain the phenomenon of lexical changes noted in ref. 1.

"Colloquialization" as a sociocultural phenomenon could have occurred for several reasons, such as popularization of mass media, population size, and competition of text readability. Biber and Finegan (7) have pointed out that popular

Author affiliations: <sup>a</sup>Department of Linguistics, University of Tübingen, 72074 Tübingen, Germany

Author contributions: K.S. analyzed data and wrote the paper.

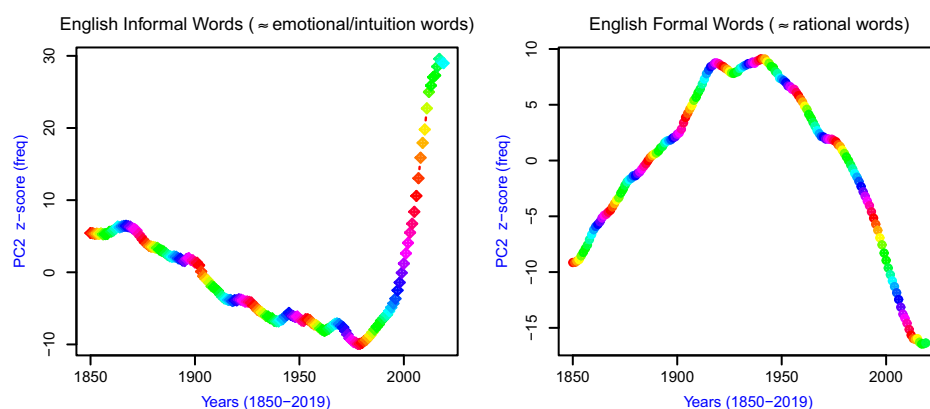
The author declares no competing interest.

Copyright © 2022 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

The datasets and programming script are available at <https://osf.io/5vbk2/>.

<sup>1</sup> Email: kun.sun@uni-tuebingen.de.

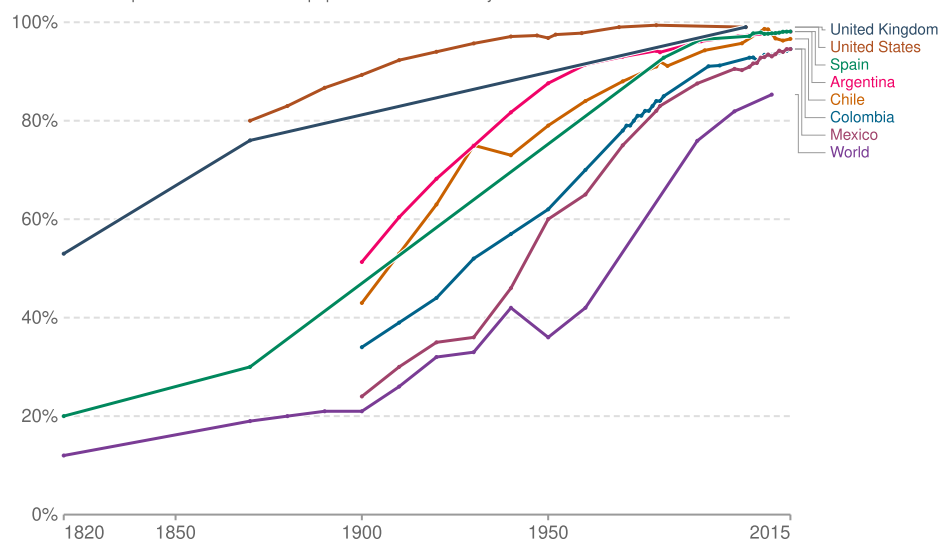
Published June 21, 2022.



**Fig. 1.** Historical changes of informal words and formal words in English. The y axis represents PC2 scaled normalized frequency. The number of either formal words or informal words is 220.

## Literacy rate, 1820 to 2015

Estimates correspond to the share of the population older than 14 years that is able to read and write.



Source: WDI, CIA World Factbook, & other sources

OurWorldInData.org/literacy • CC BY

Note: Specific definitions and measurement methodologies vary across countries and time. See the 'Sources'-tab for more details.

**Fig. 2.** Historical changes of the literacy rate across English-speaking countries, Spanish-speaking countries, and the world (1820 to 2015). The literacy rate in United Kingdom, United States, and Spain reached almost 100% in the 1980s. In the following years, there was only a very tiny increase in this indicator. The data are from <https://ourworldindata.org/literacy>. The data on other English-speaking countries are not included on this website.

literacy enhanced mass education and fostered a shift toward more oral styles (i.e., colloquialization) (also see ref. 8). Fig. 2 illustrates the changes in the literacy rate from 1850 to the present, revealing the fact that the times in

which the literacy rate changes coincide with those changes presented in ref. 1. More data and studies support the hypothesis of the effects of colloquialization and the literacy rate on language changes than that of the “post-truth era”.

1. M. Scheffer, I. van de Leemput, E. Weinans, J. Bollen, The rise and fall of rationality in language. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2107848118 (2021).
2. J.-B. Michel et al.; Google Books Team, Quantitative analysis of culture using millions of digitized books. *Science* **331**, 176–182 (2011).
3. L. Hou, D. Smith, Drivers of English syntactic change in the Canadian parliament, *Proc. SCIL* **4**, 51–60 (2021).
4. K. Hyland, F. K. Jiang, Is academic writing becoming more informal? *Engl. Specif. Purposes* **45**, 40–51 (2017).
5. T. Hiltunen, J. Räikkönen, J. Tyrkkö, Investigating colloquialization in the British parliamentary record in the late 19th and early 20th century. *Lang. Sci.* **79**, 101270 (2020).
6. C. Mair, *Twentieth-Century English: History, Variation and Standardization* (Cambridge University Press, 2006).
7. D. Biber, E. Finegan, Drift and the evolution of English style: A history of three genres. *Language* **65**, 487–517 (1989).
8. K. Sun, R. Wang, The evolutionary pattern of language in English fiction over the last two centuries: Insights from linguistic concreteness and imageability. *SAGE Open* **12**, 10.1177/21582440211069386 (2022).